| | |
|---|---|
| Title | Stream Cipher Systems Using a Chaotic Sequence of I.I.D Random Variables (5th Workshop on Stochastic Numerics) |
| Author(s) | Kohda, Tohru |
| Citation | 数理解析研究所講究録 (2001), 1240: 74-87 |
| Issue Date | 2001-12 |
| URL | http://hdl.handle.net/2433/41604 |
| Right | |
| Type | Departmental Bulletin Paper |
| Textversion | publisher |

数理解析研究所講究録 1240 巻 2001 年 74-87



# Stream Cipher Systems Using a Chaotic Sequence of I.I.D Random Variables

## i.i.d. カオス系列を用いたストリーム暗号システム

Tohru KOHDA

香田 徹

e-mail: kohda@csce.kyushu-u.ac.jp

Department of Computer Science and Communication Engineering,
Kyushu University,
Hakozaki 6-10-1, Higashi-ku, Fukuoka, 812-8581, Japan

九州大学大学院システム情報科学研究院情報工学部門
〒812-8581 福岡市東区箱崎 6 丁目 10-1

**Abstract**— A sequence of i.i.d. binary random variables has found significant applications in modern digital communication systems, such as a spreading spectrum sequence for a DS/CDMA system and a key-stream sequence for a stream cipher system. In this paper, after reviewing the generation method of sequences of i.i.d. binary random variables generated by chaotic dynamics we give a stream cipher system using a sequence of i.i.d. binary random variables which has cryptographic resistance for several kinds of cryptanalysis based on correlation functions.

**Keywords**— stream cipher, keystream sequences, random sequence, i.i.d. sequence, chaotic dynamics, Cryptanalysis, auto-/cross-correlation function, secret key

## 1 Introduction

A sequence of independent and identically distributed (*i.i.d.*) binary random variables [1]–[4] has found significant applications in modern digital communication systems such as in spread spectrum (SS) communication systems [5] or cryptosystems [6, 7] as well as in computational applications requiring random numbers [8, 9]. Such a binary sequence can be generated in various ways. Nevertheless, linear feedback shift register (LFSR) sequences which have already been thoroughly investigated based on finite field theory are employed in nearly all the methods [10].

It is, however, well known from probability theory [1]–[3] and ergodic theory [4] that coin tossing is a *theoretic model* that generates a sequence of i.i.d. binary random variables in the following sense. Imagine $n$ independent tosses of a fair coin. If we label 1 and 0 instead of head and tail, then the resulting binary sequence of length $n$ gives uniquely both the associated dyadically rational number and one of $2^n$ disjoint half-open subintervals of width

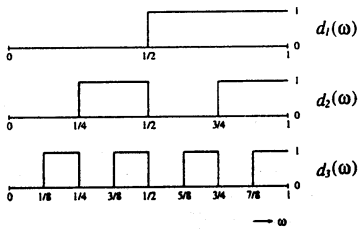



Figure 1: Binary functions $d_k(\omega)$ which define Rademacher functions $r_k(\omega)$.

$2^{-n}$ whose left endpoint is the dyadically rational number. On the other hand, choose a dyadically rational number $\omega$ *at random* and consider the 'map' (or 'transformation') on the half-open unit interval $[0,1)$, defined by $\tau(\omega) = 2\omega \,(\mathrm{mod}\, 1)$, referred to as the *dyadic map.* Then the dyadic map shifts the digit one place to the left of the binary expansion of $\omega$. In such an expansion, the binary function is sometimes referred to as the *Rademacher function.* Thus throwing a die may be seen as both a *deterministic* and a *random* event.

Ulam and von Neumann [11] pointed out the logistic map, the most famous chaotic one, is a good candidate of pseudorandom number generators. On the other hand, several encryption algorithms based on chaos theory [12]–[18] have been discussed (see, for example, the survey in [18]).

We have recently given several methods for generating a binary sequence based on chaotic dynamics [19]–[21] and a simple sufficient condition for some class of ergodic maps to produce a sequence of i.i.d. binary random variables [23, 24]. Furthermore, we have proposed a cryptosystem in a floating point environment whose keystream sequences are chaotic bit sequences [25, 26].

In this paper, we first review a sufficient condition [23, 24] for some class of ergodic maps and their associated binary functions to produce a sequence of i.i.d. binary random variables. Seconddly, we give a stream cipher system whose running key is a sequence of i.i.d binary random variables and show such a system has its cryptographical strength [25, 26].

## 2 Design of Binary Sequences Using PM Maps

### 2.1 Sequences Generated by Bernoulli Shift and Rademacher Function

Let us start by reviewing fundamental subjects from the textbooks of elementary probability theory [1]–[3] and ergodic theory [4].

Let $\omega$ be a point drawn at random from the half-open unit interval $[0,1)$. With each $\omega$ associate its nonterminating binary expansion

$$\omega = \sum_{k=1}^{\infty} \frac{d_k(\omega)}{2^k} = 0.d_1(\omega)d_2(\omega)\cdots, \text{ where } d_i(\omega) \in \{0,1\}. \tag{1}$$

Imagine now a coin with faces labeled 1 and 0 instead of the usual heads (H) and tails (T). If $\omega$ is drawn at random, $\{d_k(\omega)\}_{k=1}^{\infty}$ behaves as if it resulted from an infinite sequence of tosses of a fair coin On the other hand, define the map $\tau_B(\cdot)$, referred to as the *Bernoulli shift* or *dyadic map* by $\tau_B(\omega) = 2\omega$, mod 1 which gives $\tau_B(\omega) = 0.d_2(\omega)d_3(\omega)\cdots$. This implies that $\tau_B(\cdot)$ shifts the digits one place to the left, namely, $d_k(\tau_B(\omega)) = d_{k+1}(\omega)$, for $k \geq 1$. The

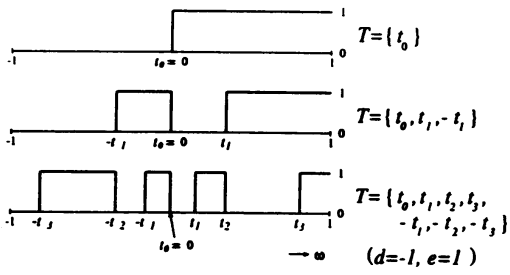


Figure 2: *Independent* pulse functions defined by $C_T(\omega)$.

functions $d_k(\omega), k = 1, 2, \cdots$ are often replaced by the *Rademacher functions* [1, 3] defined as, $r_k(\omega) = 1 - 2d_k(\omega)$, $k = 1, 2, \cdots$. Since the functions $r_k(\omega)$, $k = 1, 2, \cdots$ furnish us with a *model* of independent tosses of a "fair" coin, their values give rise to "independent random variables". Figure 1 shows the binary functions $d_k(\omega)$.

## 2.2 Binary Sequences Generated by Ergodic Maps

In the application of chaotic dynamics to cryptosystem, the dyadic map is impracticable with the help of a computer with its limited accuracy in the sense that the period of its chaotic orbit is very short. Of cruicial importance is the question whether any other maps than the dyadic map can realize coin tossing or not.

Let us start with giving three simple methods [19]–[21] for obtaining binary sequences from chaotic real-valued sequences $\{\omega_n\}_{n=0}^{\infty}$, generated by an ergodic map $\tau(\cdot) : J = [d, e] \to J$

$$\omega_{n+1} = \tau(\omega_n),\ \omega_n = \tau^n(\omega_0) \in J,\ n = 1, 2, \cdots. \tag{2}$$

*Method-1:* Define a threshold function $\Theta_t(\omega)$ as $\Theta_t(\omega) = \begin{cases} 0 & \text{for } \omega < t, \\ 1 & \text{for } \omega \geq t \end{cases}$ and its complementary function $\overline{\Theta}_t(\omega) = 1 - \Theta_t(\omega)$.

*Method-2:* For $\omega(|\omega| \leq 1)$ introduce its binary representation of:

$$|\omega| = 0.A_1(\omega)A_2(\omega)\cdots A_i(\omega)\cdots, A_i(\omega) \in \{0, 1\} \tag{3}$$

where

$$A_i(\omega) = \bigoplus_{r=1}^{2^i} \left\{\Theta_{-\frac{r}{2^i}}(\omega) \oplus \Theta_{\frac{r}{2^i}}(\omega)\right\} = \sum_{r=1}^{2^i} (-1)^{r-1} \left\{\Theta_{\frac{r}{2^i}}(\omega) + \overline{\Theta}_{-\frac{r}{2^i}}(\omega)\right\}, \tag{4}$$

and $\oplus$ denotes modulo 2 addition.

*Method-3:* [23] Define a binary function with the set of thresholds, $T = \{t_r\}_{r=0}^{M_T}$, defined by

$$C_T(\omega) = \bigoplus_{r=0}^{M_T} \Theta_{t_r}(\omega) = \sum_{r=0}^{M_T} (-1)^r \Theta_{t_r}(\omega), \tag{5}$$

then we get the following examples:

*Example 1:* $C_T(\omega) = \Theta_{t_0}(\omega)$ when $M_T = 0$.

*Example 2:* The value of $\frac{\omega - d}{e - d} \in [0, 1]$ has

$$\frac{\omega - d}{e - d} = 0.B_1(\omega)B_2(\omega)\cdots B_i(\omega)\cdots, \text{ where } B_i(\omega) \in \{0, 1\} \tag{6}$$

which implies that $C_T(\omega) = B_i(\omega)$, when $M_T = 2^{i-1}, t_\tau = (e - d)\frac{r}{2^i} + d$. If the interval $J = [0, 1]$, then $A_i(\omega) = B_i(\omega)$. The binary function $C_T(\omega)$ is referred to as the binary pulse function. Figure 2 shows its examples. Note that $B_i(\omega) = d_i(\omega)$ when $\tau(\cdot) = \tau_B(\cdot)$.

## 2.3 The Equidistributivity Property EDP and the Constant Summation Property CSP

It is an important problem whether any other maps and their associated binary functions than the dyadic map and the Rademacher function, respectively can realize a sequence of i.i.d. binary random variables or not. Recently we have replied in the affirmative: There is a wide class of ergodic maps with the equidistributivity property (EDP) and their associative binary functions with the constant summation property (CSP) [23, 24]. To review this, we begin by defining the class of maps to be discussed and introducing several definitions needed in this discussion.

Define a piecewise monotonic (PM) onto ergodic map $\tau(\cdot) : J = [d, e] \to J$ that satisfies the following conditions:

**i)** there is a partition $d = d_0 < \cdots < d_{N_\tau} = e$ of $J$ such that for each integer $i = 1, \cdots, N_\tau$, $(N_\tau \geq 2)$ the restriction of $\tau(\cdot)$ to the interval $J_i = [d_{i-1}, d_i)$, denoted by $\tau_i(\omega)$, is a $C^2$ function; as well as

**ii)** $\tau(J_i) = (d, e)$;

**iii)** $\tau$ has a unique ACI measure, denoted by $f^*(\omega)d\omega$.

**Definition 1** *For PM onto map with its ACI measure, the partition $\{J_i\}_{i=1}^{N_\tau}$ is referred to as the trivial partition of the interval $J$.*

**Definition 2** *[27, 28] The Perron-Frobenius operator (or 'transfer' operator) $P_\tau$ acting on the function of bounded variation $H(\omega) \in L^\infty$ for $\tau(\omega)$ is defined as*

$$P_\tau H(\omega) = \frac{d}{d\omega}\int_{\tau^{-1}([d,\omega])} H(y)dy = \sum_{i=1}^{N_\tau} |g_i'(\omega)|H(g_i(\omega)), \tag{7}$$

*where $g_i(\omega) = \tau_i^{-1}(\omega)$ is the $i-th$ preimage of $\omega$.*

This implies $P_\tau$ is the (discrete-time) evolution operator of the probability density function under the map $\tau(\cdot)$ [27]. Note that the ACI measure $f^*(\omega)d\omega$ satisfiies the *invariant* equation $P_\tau f^*(\omega) = f^*(\omega)$, referred to as the P-F equation.

**Definition 3** *[23] The map is said to satisfy equidistributivity property (EDP) if the relation*

$$|g_i'(\omega)|f^*(g_i(\omega)) = \frac{f^*(\omega)}{N_\tau}, \; 1 \leq i \leq N_\tau \tag{8}$$

**Definition 4** *[23] If for a class of maps with EDP its associated function $F(\cdot)$ satisfies*

$$\frac{1}{N_\tau}\sum_{i=1}^{N_\tau} F(g_i(\omega)) = \mathbf{E}[F], \tag{9}$$

*then $F(\cdot)$ is said to satisfy the constant summation property (CSP), where $\mathbf{E}[F]$, the ensemble average of $F(\omega)$, is defined in Definition 8.*

**Remark 1** *[23] Consider a random variable $F(\omega)$ and its weighted random function $F(\omega)f^*(\omega)$ with the weight function $f^*(\omega)$. Observe*

$$P_\tau\{F(\omega)f^*(\omega)\} = \frac{1}{N_\tau}\sum_{i=1}^{N_\tau} |g_i'(\omega)| f^*(g_i(\omega))F(g_i(\omega)). \tag{10}$$

*An easy consequence of the definition is that if given a weighted density function with respect to a random variable $F(\omega)$ with $\mathbf{E}[F] \neq 0$, defined by $F(\omega)f^*(\omega)$, both of the CSP of $F(\omega)$ and the EDP of $f^*(\omega)$ are satisfied, then $\dfrac{F(\omega)}{\mathbf{E}[F]}f^*(\omega)$ is an invariant function under the map $\tau(\cdot)$.*

**Definition 5** *(Topological Conjugation) [28, 31] Two transformations $\overline{\tau}: \overline{I} \to \overline{I}$ and $\tau: I \to I$ on intervals $\overline{I}$ and $I$ are called topological conjugate if there exists a homeomorphism $h$: $\overline{I} \xrightarrow{\text{onto}} I$, such that $\tau(\omega) = h \circ \overline{\tau} \circ h^{-1}(\omega)$.*

Suppose $\tau(\cdot)$ and $\overline{\tau}(\cdot)$ have their ACI measures and denote them by $f^*(\omega)d\omega$ and $\overline{f}^*(\overline{\omega})d\overline{\omega}$, respectively. Then, under the *topological conjugation*, these ACI measures have the relation

$$f^*(\omega) = \left|\frac{dh^{-1}(\omega)}{d\omega}\right| \overline{f}^*(h^{-1}(\omega)). \tag{11}$$

It is easily checked from simple calculation that the EDP is invariant under the topological conjugation.

Examples of the maps with their ACI measures which satisfy the EDP are listed as follows.

(1) $R$-adic map ($N_\tau = R, R = 2, 3, 4, \cdots$)

$$\tau_R(\omega) = R\omega \bmod 1,\ \omega \in [0,1],\ f^*(\omega)d\omega = d\omega,$$

(2) logistic map ($N_\tau = 2$) [11]

$$L_2(\omega) = 4\omega(1-\omega), \omega \in [0,1],\ f^*(\omega)d\omega = \frac{d\omega}{\pi\sqrt{\omega(1-\omega)}},$$

(3) Chebyshev map of degree $p$ ($N_\tau = p$, $p = 2, 3, 4, \cdots$) [29]–[31]

$$T_p(\omega) = \cos(p\cos^{-1}\omega),\ \omega \in [-1,1],\ f^*(\omega)d\omega = \frac{d\omega}{\pi\sqrt{1-\omega^2}},$$

(4) Zigzag map with $p$ branches ($N_\tau = p$, $p = 2, 3, 4, \cdots$) [31]

$$N_p(\omega) = (-1)^{\lfloor p\omega \rfloor} p\omega,\ \omega \in [0,1],\ f^*(\omega)d\omega = d\omega,$$

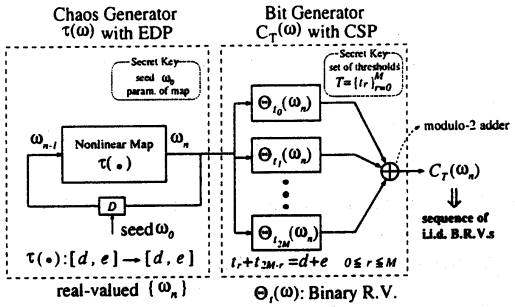



Figure 3: Sequence generator of i.i.d. binary random variables for a stream cipher system using chaotic dynamics.

where $\lfloor x \rfloor$ denotes the largest integer not greater than $x$ and $N_2(\omega)$ is referred to as the tent map.

Ulam and von Neumann [11] pointed out that the logistic map $L_2(\omega)$ and the tent map $N_2(\overline{\omega})$ are topological conjugate via $h^{-1}(\omega) = \frac{2}{\pi}\sin^{-1}\sqrt{\omega}$. On the other hand, Grossmann and Thomae [31] observed that $T_p(\omega)$ and $N_p(\overline{\omega})$ are topological conjugate via $h(\overline{\omega}) = \cos \pi\overline{\omega}$.

Now let us consider a stationary real-valued sequences $\{F(\omega_n)\}_{n=0}^{\infty}$, where $\omega_n = \tau^n(\omega_0)$.

**Definition 6** *(Birchoff individual ergodic theorem) [4, 28] The time average of any $L_1$ function $F(\cdot)$ along a chaotic real-valued trajectory $\{\omega_n\}_{n=0}^{\infty}$ generated by the map (2), denoted by $Average_T\{F\}$, defined by $Average_T\{F\} = \lim_{T\to\infty}\frac{1}{T}\sum_{n=0}^{T-1} F(\omega_n)$ is equal almost everywhere to the expectation of $F(\omega)$, defined by $\mathbf{E}[F(X_n)] = \int_I F(X_n)f^*(\omega)d\omega$. From the stationarity of process, we denote $\mathbf{E}[F(X_n)]$ by $\mathbf{E}[F(X)]$ or simply $\mathbf{E}[F]$* [1] .

Of major importance to the investigation of statistical properties of two sequences $\{F(\tau^n(\omega))\}_{n=0}^{\infty}$ and $\{G(\tau^n(\omega))\}_{n=0}^{\infty}$ is their 2nd-order cross-covariance function with delay $\ell = 0, 1, \cdots$ which is defined by

$$\rho(\ell; F, G) = \int_I (F(\omega) - \mathbf{E}[F])(G(\tau^\ell(\omega)) - \mathbf{E}[G])f^*(\omega)d\omega, \tag{12}$$

which denotes the auto-covariance function when $F = G$.

**Remark 2** *Note that the opertaor $P_\tau(\cdot)$ has the most important property of the P-F operator $P_\tau$:*

$$\int_J F(\omega)G(\tau(\omega))d\omega = \int_J P_\tau\{F(\omega)\}G(\omega)d\omega, \tag{13}$$

*where $G(\tau(\omega))$ in the integrand in the lhs is referred to as the Koopman operator, as the adjoint operator of the Perron-Frobenius operator in a Hilbert space [28, 32]. Using (13), we get the rewritten form of (12)*

$$\rho(\ell; F, G) = \int_I P_\tau^\ell\{F(\omega)f^*(\omega)\}G(\omega)d\omega - \mathbf{E}[F]\mathbf{E}[G]. \tag{14}$$

[1] It should be noted that expectation $\mathbf{E}[\cdot]$ is used here in a sense applicable to deterministic sequences of the form (2), and differs from the standard one for stochastic sequences which will involve a joint probability density function.

Equation (10) and condition ii) of the onto map $\tau(\cdot)$ give

**Lemma 1** *[23] The PM onto maps $\tau(\cdot)$ with EDP satisfy*

$$P_\tau\{\Theta_{d_i}(\omega)f^*(\omega)\} = \mathbf{E}[\Theta_{d_i}]f^*(\omega). \tag{15}$$

**Lemma 2** *[23] For a class of maps with EDP, the following three symmetric properties:*

**S1:** *the symmetric binary function $C_T(\omega)$, defined as*
$t_r + t_{2M-r} = d + e$, $r = 0, 1, \cdots, M$;

**S2:** *the symmetric ACI measure, defined as*
$f^*(d + e - \omega) = f^*(\omega)$;

**S3:** *the symmetric map, defined as*
$\tau(d + e - \omega) = \tau(\omega), \quad \omega \in I$

*give*

$$P_\tau\{C_T(\omega)f^*(\omega)\} = \mathbf{E}[C_T]f^*(\omega). \tag{16}$$

**Remark 3** *Relation (16) is a generalized version of (15); the CSP of $F(\omega)$ guarantees their zero correlations, $\rho(\ell, F, G) = 0$, for 'positive' delay $\ell$ irrespective of $G(\omega)$.*

Next, let $\mathbf{U} = U_0U_1\cdots U_{m-1}$ be an arbitrary string of $m$ binary digits where $U_n \in \{0,1\}$ $(0 \le n \le m-1)$. Then there are $2^m$ possible strings. Let $\mathbf{u}^{(r)} = u_0^{(r)}u_1^{(r)}\cdots u_{m-1}^{(r)}$ be the $r$-th string. Introducing a binary random variable

$$\Gamma_n(\omega; F, \mathbf{u}^{(r)}) = F(\omega)u_n^{(r)} + \overline{F}(\omega)\overline{u}_n^{(r)} \tag{17}$$

for any binary function $F(\omega)$ with CSP, we can get [23]

*Theorem 1: The probability of the event $\mathbf{u}^{(r)}$ in an infinite binary sequence $\{C_T(\omega_n)\}_{n=0}^{\infty}$ is given by*

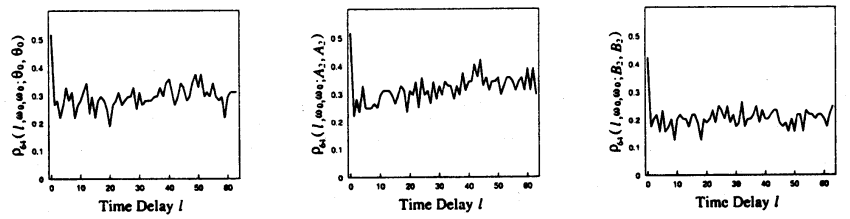
$$\begin{aligned} Pr(\mathbf{u}^{(r)}; C_T) &= \int_I \left\{\prod_{n=0}^{m-1} \Gamma_n(\omega_n; C_T, \mathbf{u}^{(r)})\right\} f^*(\omega)d\omega \\ &= \mathbf{E}[C_T]^s(1 - \mathbf{E}[C_T])^{m-s}, \end{aligned} \tag{18}$$

*where $s$ is the number of 1 in $\{u_n^{(r)}\}_{n=0}^{m-1}$.*
This implies $\{C_T(\tau^n(\omega))\}_{n=0}^{\infty}$ is a sequence of i.i.d. binary random variables with probability $\mathbf{E}[C_T]$. Note that we can get a fair Bernoulli sequence when $\mathbf{E}[C_T] = \frac{1}{2}$, that is, an $m$-distributed binary random sequence. Figure 2 shows an example of symmetric binary function $C_T(\omega)$.

# 3 Stream Cipher System

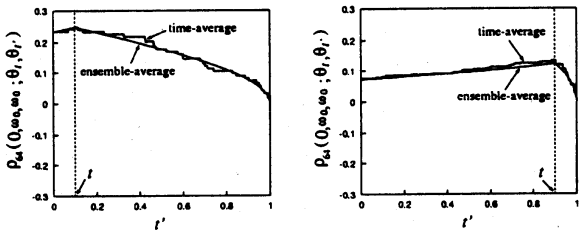
## 3.1 Stream Cipher System Using Chaotic Dynamics

We employ the so-called binary additive stream ciphers in which a short secret key $\vec{K} = (s_1.s_2, \cdots, s_K)$ is used only to control a keystream generator. We have secret key parameters




(a) $\rho_{64}(0, \omega_0, \omega_0; \Theta_0, \Theta_0)$ (b) $\rho_{64}(\ell, \omega_0, \omega_0; A_2, A_2)$ (c) $\rho_{64}(\ell, \omega_0, \omega_0; B_2, B_2)$

Figure 4: Autocorrelation functions




$\rho_{64}(0, \omega_0, \omega_0; \Theta_t, \Theta_{t'})$ $\rho_{64}(0, \omega_0, \omega_0; \Theta_t, \Theta_{t'})$

Figure 5: Crosscovariance functions

$\omega_0$ and $T$ as well as parameters of the map $\tau(\cdot)$. As a nonlinear ergodic map $\tau(\cdot)$, we use the Chebyshev maps of degree $k$, defined by

$$\tau(\omega) = \cos(k \cos^{-1} \omega), \quad k = 2, 3, 4, \cdots \tag{19}$$

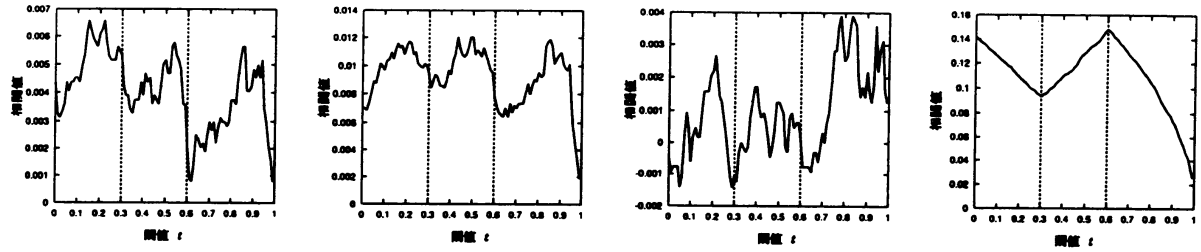
in the following numerical experiments. The central problem in stream cipher cryptography is the difficulty of efficiently generating long running-key sequences from a short and random key. It is obvious that a sequence of i.i.d. binary random variables is one of good candidates of running-key sequences. Such a situation motivated us to propose a stream cipher system [25, 26] whose running-key sequence is a sequence of *i.i.d.* binary random variables based on chaotic dynamics, *e.g., symmetric binary sequence* $\{C_T(\omega_n)\}_{n=0}^{\infty}$, defined by (5) with $M_T = 2M$ and the symmetric threshold set $T = \{t_r\}_{r=0}^{2M}$, defined by condition S1 in *Lemma* 2. Figure 3 shows a stream cipher system based on $C_T(\omega)$.

## 3.2 Statistical Properties of Chaotic Bit Sequences

The set of thresholds $T = \{t_r\}_{r=0}^{2M}$ of $C_T(\cdot)$ are candidates of secret key parameters. However, such parameters are not cryptographically secure and should not be used because the statistics of chaotic bit sequences are not sensitive to $\{t_r\}$ as follows [5] .

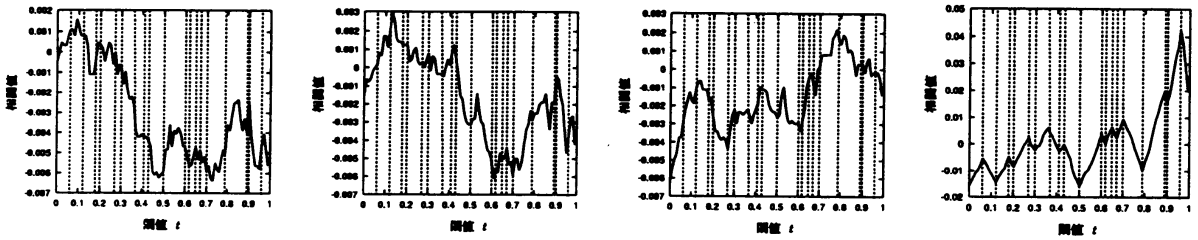
Consider the time-averaged crosscorrelation function between sequences $\{F(\omega_n)\}_{n=0}^{\infty}$ with

[5] This situation is similar to the one that chaos synchronization systems are insensitive to their system parameters.



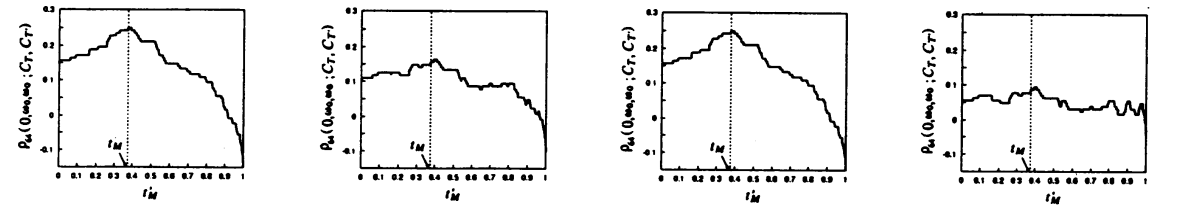
(a) 7-bit precision (b) 8-bit precision (c) 9-bit precision (d) 10-bit precision

Figure 6: Cryptanalysis of $t$ of $C_T$ sequence using $\rho_{64}(0, \omega_0, \omega_0'; C_T, \Theta_t)$ $(M = 2)$.



(a) 7-bit precision (b) 8-bit precision (c) 9-bit precision (d) 10-bit precision

Figure 7: Cryptanalysis of $t$ in $C_T$ sequence using $\rho_{64}(0, \omega_0, \omega_0'; C_T, \Theta_t)$ $(M = 19)$.



(a) $M = 3$, $\epsilon = 0.01$ (b) $M = 20$, $\epsilon = 0.01$ (c) $M = 3$, $\epsilon = 0.02$ (d) $M = 20$, $\epsilon = 0.02$

Figure 8: Cryptanalysis of $t_M$ using $\rho_{64}(0, \omega_0, \omega_0; C_T, C_{T'})$.

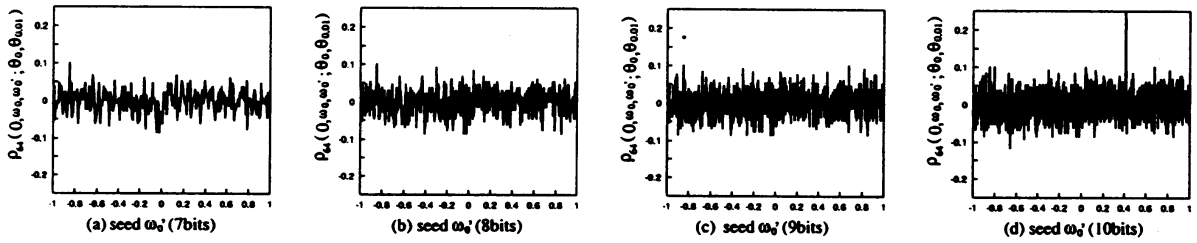



Figure 9: Cryptanalysis using $\rho_{64}(0, \omega_0, \omega_0'; \Theta_t, \Theta_{t'})$.

an initial seed $\omega_0$ and $\{G(\omega_n')\}_{n=0}^{\infty}$ with another initial seed $\omega_0'$, defined by

$$\rho_N(\ell, \omega_0, \omega_0'; F, G) = \frac{1}{N}\sum_{n=0}^{N-1} F(\omega_n)G(\omega_{n+\ell}'), \tag{20}$$

where the subscripts are taken mod $N$. Note that $r_N(\ell, \omega_0, \omega_0'; \Theta_t, \Theta_{t'})$ has a large value only at $\ell = 0$ because of its no-correlation property. Figure 4 shows autocorrelation functions with their no-correlation property. On the other hand, the fluctuated and smooth curves in Figure 5 indicate respectively the time-average $r_{64}(0, \omega_0, \omega_0; \Theta_t, \Theta_{t'})$ of cross-correlation between $\Theta_t(\cdot)$ and $\Theta_{t'}(\cdot)$ versus $t'$, and the ensemble-averaged one $\tilde{\rho}(0; \Theta_t, \Theta_{t'})$, where $\tilde{\rho}(\ell; F, G)$ is defined by

$$\tilde{\rho}(\ell; F, G) = \int_J F(\omega)G(\tau^{\ell}(\omega))f^*(\omega)\, d\omega. \tag{21}$$

Observe that this figure implies that the estimation of the parameter $t$ is easy if the seed $\omega_0$ and parameters of the map are previously known and hence that the key $t$ is not cryptographically secure and should not be used. Thus the set of thresholds, $T = \{t_r\}_{r=0}^{2M}$, may be used only to satisfy the i.i.d. condition of sequences.

## 3.3 Cryptanalysis

### 3.3.1 Cryptanalysis of thresholds $t_i$

Figure 6 and 7 show searches of the thresholds $\{t_r\}$ using cross-covariance functions $\rho_{64}(0, \omega_0, \omega_0'; C_T, \Theta_t)$, where $T = \{0.3.0.6\}$ and $T = \{0.06, 0.12, 0.18, 0.2, 0.27, 0.3, 0.36, 0.41, 0.43, 0.5, 0.6, 0.62, 0.65, 0.67, 0.7, 0.79, 0.89, 0.9, 0.96\}(M = 19)$, respectively and where $\omega_0'$ is nearly equal to $\omega$ within (a) 7-bit precision, (b) 8-bit precision, (c) 9-bit precision, and (d) 10-bit precision. Observe that these figures imply that the estimation of the parameter $t$ and hence the estimation of the parameter $T = \{t_r\}_{r=0}^{2M}$ is easy if the seed $\omega_0$ and parameters of the map are previously known. Unless $\omega_0'$ is equal to $\omega$, however, the search of $t$ is not easy. Eventhough, the key $t$ is not cryptographically secure and should not be used. On the contrary, the parameters $T$ may be used only to satisfy the conditions for a sequence to be one of *i.i.d.* binary random variables. Figure 8 shows $\rho_{64}(0, \omega_0, \omega_0; C_T, C_{T'})$ versus $t_M'$ to search $t_M$, where $T = \{t_r\}_{r=1}^{M}, T' = \{t_r'\}_{r=1}^{M}$, and $\{t_r \pm \epsilon\}_{r=1}^{M-1} = \{t_r'\}_{r=1}^{M-1}$.

### 3.3.2 Cryptanalysis of seed $\omega_0$

On the other hand, a seed $\omega_0$ (*i.e., an initial condition*) is the best choice of a secret key in stream cipher cryptography because of the *SDIC* (sensitive dependence on initial conditions) property and the ergodicity of chaotic maps as discussed in Section 2. Figures 9 shows an illustration of cryptanalysis using the statistics $r_{64}(0, \omega_0, \omega_0'; \theta_t, \theta_{t'})$, $r_{64}(0, \omega_0, \omega_0'; C_T, C_{T'})$, and $\rho_{64}(0, \omega_0, \omega_0'; C_T, C_T)$, respectively to search a 10-digit seed $\omega_0 = (0.0110100101)_2 \simeq 0.4113281\ldots$, a partial secret key of a bit sequence $\{C_T(\omega_n)\}_{n=0}^{\infty}$ provided that all of other parameters are known within the limited accuracy, ($\omega_0'$ is nearly equal to $\omega$ within (a) 7-bit precision, (b) 8-bit precision, (c) 9-bit precision, and (d) 10-bit precision). In this search all possible $2^{10}$ $\omega_0'$'s are scanned. Figure 10 shows cryptanalysis of $t_M$ and $\omega_0$ using $\rho_{64}(0, \omega_0, \omega_0'; C_T, C_{T'})$, where $T = \{t_r\}_{r=1}^{M}$, $T' = \{t_r'\}_{r=1}^{M-1}$, and $\{t_r\}_{r=1}^{M-2} = \{t_r'\}_{r=1}^{M-2}$. Figure 11 shows cryptanalysis of $\omega_0$ using $r_{64}(0, \omega_0, \omega_0'; C_T, C_T)$. We can find that we need exhaustive searches of $\omega_0'$ even if the parameters of the map, *e.g.*, degree $p$ of $T_p(\omega)$ and the set of thresholds are previously known because $r_{64}(0, \omega_0, \omega_0'; C_T, C_T)$ has a peak only when

$\omega_0'$ is equal to $\omega_0$ completely in a given precision [26]. This implies that this strategy is computationally infeasible because of the large key space of $\omega_0$.

### 3.4 Hierchical Structure

A chaotic bit sequence $\{C_T(\omega_n)\}_{n=0}^{\infty}$ is a sequence of *i.i.d.* binary random variables under certain conditions[23]. The *i.i.d.* property make it easier to construct chaotic bit generators with hierarchical structures which can also produce sequences of *i.i.d.* binary random variables. The first one, called *level 1 hierarchy*, is based on the no-correlation between any two sequences, $\{C_T(\omega_n)\}_{n=0}^{\infty}$ and $\{C_{T'}(\omega_{n+\ell})\}_{n=0}^{\infty}$, from a seed for $\ell \geq 1$. The no-correlation property between $\{C_T(\omega_n)\}_{n=0}^{\infty}$ and $\{C_{T'}(\omega_n')\}_{n=0}^{\infty}$ is defined by[23]

$$\rho_N(\ell, \omega_0, \omega_0'; C_T, C_{T'}) - \frac{1}{4} \to \left(Q_{TT'} - \frac{1}{4}\right)\delta(\ell)\delta(\omega_0 - \omega_0'), \quad \text{as } N \to \infty \tag{22}$$

where $Q_{TT'}$ denotes a nontrivial value which is a function of $T$ and $T'$. The second one, called *level 2 hierarchy*, is based on the no-correlation between any two sequences from different seeds. Such a sequence $\{C_T(\omega_n)\}_{n=0}^{\infty}$ is referred to as *level 0 hierarchy*.

In order to increase the number of different secret keys, we can use the modulo 2 addition of chaotic symmetric binary sequences $\{C_{T_j}(\omega_n)\}_{n=0}^{\infty}$ with a symmetric threshold set

$$T_j = \{t_i(j)\}_{i=0}^{2M(j)}, \quad 0 \leq j \leq m-1. \tag{23}$$

For a positive integer $d_i$ $(1 \leq i \leq m-1)$, we get

$$D_{\vec{T},\vec{d}}(\omega_n) = C_{T_0}(\omega_n) \oplus C_{T_1}(\omega_{n-d_1}) \oplus \cdots \oplus C_{T_{m-1}}(\omega_{n-d_1-\cdots-d_{m-1}}), \tag{24}$$

$$\vec{T} = \{T_0, T_1, \cdots, T_{m-1}\}, \quad \vec{d} = \{d_1, d_2, \cdots, d_{m-1}\} \tag{25}$$

which is obtained from a chaotic real-valued trajectory $\{\omega_n\}_{n=0}^{\infty}$ as shown in Figure 12 (a). Note that $\{D_{\vec{T},\vec{d}}(\omega)\}_{n=0}^{\infty}$ is also a sequence of *i.i.d.* binary random variables. Hence we call such a sequence $\{D_{\vec{T},\vec{d}}(\omega)\}_{n=0}^{\infty}$ *a generalized version of chaotic symmetric binary sequences.* For such a sequence, we can use not only $\vec{T}$ but also $\vec{d}$ as secret keys.

For $L$ seeds which are chosen statistically independently and for $L$ ergodic maps, we define

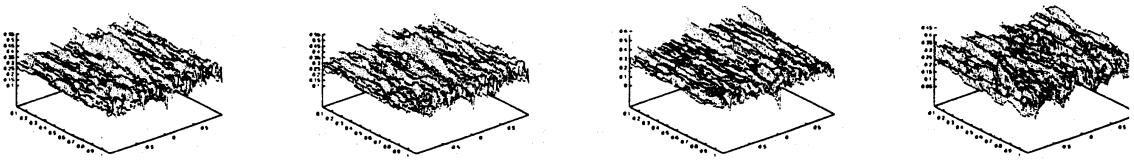
$$E_{\vec{\tau}}(\vec{\omega_n}) = D_{\vec{T}(1),\vec{d}(1)}(\omega_{n,1}) \oplus D_{\vec{T}(2),\vec{d}(2)}(\omega_{n,2}) \oplus \cdots \oplus D_{\vec{T}(L),\vec{d}(L)}(\omega_{n,L}) \tag{26}$$

$$\vec{\tau} = \{\tau_1, \tau_2, \cdots, \tau_L\}, \quad \vec{\omega_n} = \{\omega_{n,1}, \omega_{n,2}, \cdots, \omega_{n,L}\}, \ \omega_{n,s} = \tau_s^n(\omega_{0,s}) \tag{27}$$

The additive property of *i.i.d.* binary random variables allows us to get a sequence of *i.i.d.* binary random variables $\{E_{\vec{\tau}}(\vec{\omega}_n)\}_{n=0}^{\infty}$ as shown in Figure 12. Note that periods of such sequences can be longer than ones of each real-valued sequences $\{\omega_{n,s}\}_{n=0}^{\infty}$ in a finite precision computation system. Figure 12 shows (a) a bit generator $D_{\vec{T},\vec{d}}(\omega)$ and (b) a bit generator $E_{\vec{\tau}}(\vec{\omega}_n)$, respectively.

## 4 Concluding Remarks

Nevertheless, it is hard to prove that such a system implemented in a floating-point environment is *cryptographically secure* from the cryptographic point of view: *e.g.*, the period of a running sequence generated by this cryptosystem is short. Some cryptographic problems can be settled only if progress in some number-theoretic problems and/or information theoretic ones can be made. Lastly note that most of the existing chaos cryptosystems have been incapable of fully utilizing the *sensitive dependence on initial conditions property*, *i.e.*, $\omega_0$, primarily because they are based on analog circuits.



(a) $M = 3$ (b) $M = 5$ (c) $M = 10$ (d) $M = 20$

Figure 10: Cryptanalysis using $\rho_{64}(0, \omega_0, \omega_0'; C_T, C_{T'})$.

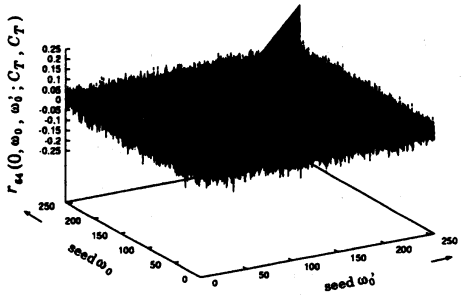
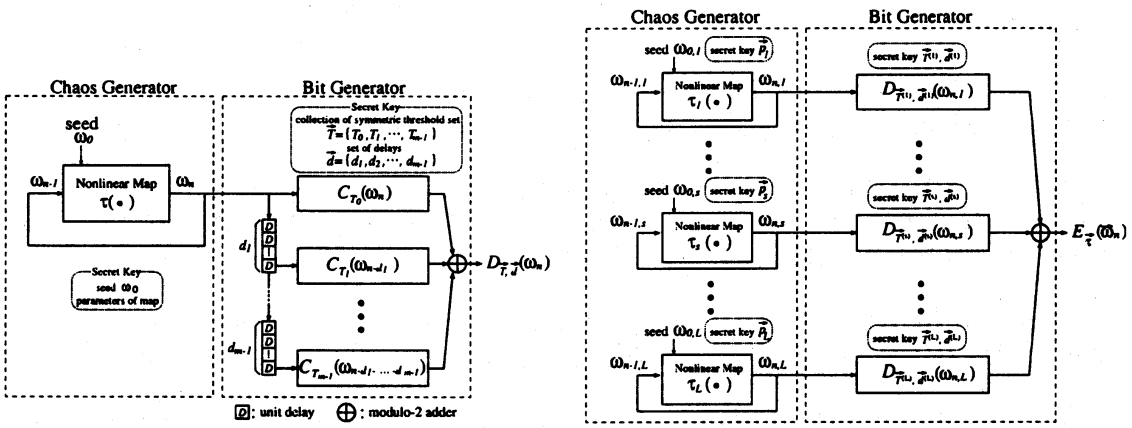



Figure 11: Cryptanalysis of $\omega_0$ using $r_{64}(0, \omega_0, \omega_0'; C_T, C_T)$.




(a) $D$–sequence generator (b) $E$–sequence generator

Figure 12: Hierachical sequence generator

## References


[1] M. Kac, "Statistical Independence in Probability Analysis and Number Theory," The Mathematical Association of America, 1959.

[2] M. Loève, "Probability Theory I," Graduate Texts in Mathematics 45, Springer-Verlag, 1977.

[3] P. Billingsley, "Probability and Measure," John Wiley & Sons, 1995.

[4] P.R. Halmos, "Lectures on Ergodic Theory," The Mathematical Society Japan, 1956.

[5] A. J. Viterbi and J. K. Omura, "Principles of Digital Communication and Coding," McGraw-Hill Book Company, Tokyo, 1979.

[6] J. L. Massey, "An introduction to contemporary cryptology," *Proc. IEEE*, **76**, no.5, 533–549, May 1988.

[7] T.W. Cusick, C.Ding, and A.Renvall, "*Stream Ciphers and Number Theory*," North Holland, 1998.

[8] J. von Neumann, Summary written by G. E. Forsythe, "Various techniques used in connection with random digit", *National Bureau of Standards, Applied Math. Series*, **12**, 36–38, 1951.

[9] D. Knuth, "The Art of Computer Programming 2, Seminumerical Algorithms," 2nd ed., Addison-Wesley, Reading, Mass, 1981.

[10] D. V. Sarwate and M. B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE*, **68**, no.3, 593–619, 1980.

[11] S. M. Ulam and J.von Neumann, "On combination of stochastic and deterministic processes", *Bull. Math. Soc.*. **53**, pp.1120,1947.

[12] R.A. Matthews, "On the derivation of a "chaotic" encryption algorithm," *Cryptologia*, **13**,pp.29-42, 1989.

[13] D.D. Wheeler and R.A. Matthews, "Supercomputer Investigations of a chaotic encryption algorithm," *Cryptologia*, **15**,pp.140–152, April, 1991.

[14] T.Habutsu, Y.Nishio, I.Sasase and S.Mori, "A secrete key cryptosystem by iterating a chaotic map," in *Proc. Advances in Cryptology*–EUROCRYPT'91. Berelin, Germany:Springer-Verlag,1991, pp.127-140.

[15] M. Götz, K. Kelber, and W. Schwarz, "Discrete-time chaotic encryption systems part I : statistical design approach" *IEEE Trans. Circuit Syst.*, **CAS-44**, no.10, 963-970, 1997.

[16] F. Dachselt, K. Kelber, and W. Schwarz, "Discrete-time chaotic encryption systems part III: cryptographical analysis",*IEEE Trans. Circuit Syst.*, **CAS-45**, no.9, pp.983–988, 1998.

[17] M.S. Bapista, "Cryptography with chaos," *Phys.Lett.A*, **240**, pp.5–54. 1998.

[18] G. Jakimoski and L. Kocarev, "Chaos and cryptography:block encryption ciphers based on chaotic maps," *IEEE Trans. Circuit Syst.*, **CAS-48**, no.2, pp.163–169, 2001.

[19] T. Kohda, A. Tsuneda, and T. Sakae, "Chaotic Binary Sequences by Chebychev Maps and Their Correlation Properties",*Proc. of the IEEE Second Int. Sympo. on Spread Spectrum Techniques and Applications*, 63–66, 1992.

[20] T. Kohda and A. Tsuneda, "Pseudonoise sequences by chaotic nonlinear maps and their correlation properties," *IEICE Trans. Communications*, **E76-B**, no.8, 855–862, 1993.

[21] T. Kohda and A. Tsuneda, "Explicit evaluations of correlation functions of Chebyshev binary and bit sequences based on Perron-Frobenius operator," *IEICE., Trans. Fundamentals*, **E77-A**, no.11, 1794–1800, 1994.

[22] D.S. Broomhead, J.P. Huke, and M.R. Muldoon, "Codes for spread spectrum applications generated using chaotic dynamical systems. *Dynamics and Stability of Systems*,**14**, pp.95–105,1999.

[23] T. Kohda and A. Tsuneda, "Statistics of chaotic binary sequences", *IEEE Trans. Information Theory*, **43**, no.1, 104–112, 1997.

[24] T. Kohda, "Sequences of i.i.d. binary random variables using chaotic dynamics", *Sequences and Their Applications*, eds. by C. Ding, T. Helleseth, and H. Niederreiter, 297–307, Springer-Verlag, 1999.

[25] T.Kohda and A.Tsuneda, "Chaotic Bit Sequences for Stream Cipher Cryptography and Their Correlation Functions", Proc. SPIE's Photonics East '95 Sympo. (Chaotic Circuit for Communication), SPIE Vol.2612, pp.86–97, 1995.

[26] T. Kohda and A. Tsuneda, "Stream cipher systems based on chaotic binary sequences",*Proc. the 19996 Symp. on Cryptography and Information Security*, SCIS96-11B, pp.1–8, 1996.

[27] O. W. Rechard, "Invariant measures for many-one transformation," *Duke Math.J.*, **23** pp.477-488, 1956.

[28] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise*, Springer-Verlag, 1994.

[29] R. L. Adler and T. J. Rivlin, "Ergodic and mixing properties of Chebyshev polynomials," *Proc. Amer. Math. Soc.*, **15**, 794, 1964.

[30] T. J. Rivlin, "Chebyshev Polynomials", John Wiley & Sons, Inc., 1990.

[31] S. Grossmann and S. Thomae, "Invariant distributions and stationary correlation functions of one-dimensional discrete processes," *Z. Naturforsch.*, **32a**, 1353–1363, 1977.

[32] B.O. Koopman, "Hamiltonian systems and transformations in Hilbert space", *Proc.Natl.Acad.Sci.*, U.S.A.,**17**,pp.315-318,1931.